



419 OIPE

RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/804,481

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

ENTERED

```

4 <110> APPLICANT: de Graaf, David
5      Lander, Eric S.
7 <120> TITLE OF INVENTION: Novel Small Nuclear RNA Vectors and Uses
8      Therefor
10 <130> FILE REFERENCE: 2825.1023-001
12 <140> CURRENT APPLICATION NUMBER: 09/804,481
13 <141> CURRENT FILING DATE: 2001-03-12
15 <150> PRIOR APPLICATION NUMBER: 60/188,304
16 <151> PRIOR FILING DATE: 2000-03-10
18 <160> NUMBER OF SEQ ID NOS: 11
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 4639
24 <212> TYPE: DNA
25 <213> ORGANISM: Homo sapiens
27 <400> SEQUENCE: 1
28 gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
29 ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
30 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgagct cggtaccccg 180
31 ggagatccgg taaggaccag cttcttttgg agagaacaga cgcagggggc ggagggaaaa 240
32 agggagaggg agacgtcact tccccttggc ggctctggca gcagatttgt cggttgagtg 300
33 gcagaaaggg agacggggac tgggcaaggg actgtcgggtg acatcacgga cagggcgact 360
34 tctatgtaga tgaggcagcg cagaggctgc tgcttcgcca cttgctgctt caccacgaag 420
35 gagttcccggt gccctgggag cgggttcagg accgctgate ggaagtgaga atcccagctg 480
36 tgtgtcaggg ctggaaaggg ctcgggagtg cgcggggcaa gtgaccgtgt gtgtaaagag 540
37 tgaggcgtag gaggtgtgtt cggggcagag gcccaagatc tcaagggccc ataacatgtg 600
38 taccatcgat tgcaggggag ataccatgat cacgaagggt gttttcccag ggcgaggctt 660
39 atccattgca ctccggatgt gctgacccct gcgatttccc caaagcttgg aaactcgact 720
40 gcataatttg tggtagtggg ggactgcgtt cgcgctttcc cctgactttc tggagtttca 780
41 aaagtagact gtacgctaac cggatcctct agagtcgacc tgcaggcatg cagaagacaa 840
42 ttagcaggca tgctggggat gcggtgggct ctatggcttc tgaggcggaa agaaccagct 900
43 ggggctctag ggggtatccc cacgcgccct gtagcggcgc attaagcgcg gcgggtgttg 960
44 tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcgccgt cctttcgctt 1020
45 tcttcccttc ctttctcgcc acgttcgccc gctttccccg tcaagctcta aatcggggca 1080
46 tccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 1140
47 gtgatggttc acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg 1200
48 agtccacggt ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct 1260
49 cggctctattc ttttgattta taagggattt tggggatttc ggcctatttg ttaaaaaatg 1320
50 agctgattta acaaaaattt aacgcgaatt aattctgttg aatgtgtgtc agttagggtg 1380
51 tggaaagtcc ccaggctccc caggcaggca gaagtatgca aagcatgcat ctcaattagt 1440
52 cagcaaccag gtgtggaaag tccccaggct ccccagcagg cagaagtatg caaagcatgc 1500
53 atctcaatta gtcagcaacc atagtcccgc ccctaactcc gcccatcccg cccctaactc 1560
54 cgcccagttc cgcccattct cgcgcccatg gctgactaat tttttttatt tatgcagagg 1620

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,481

DATE: 01/27/2002

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

```

55 ccgaggccgc ctctgcctct gagctattcc agaagtagtg aggaggcttt tttggaggcc 1680
56 taggcttttg caaaaagctc ccgggagctt gtatatccat tttcggatct gatcagcacg 1740
57 tgttgacaat taatcatcgg catagtatat cggcatagta taatacgaca aggtgaggaa 1800
58 ctaaaccatg gccaaagtga ccagtgccgt tccggtgctc accgcgcgcg acgtcgccgg 1860
59 agcggtcgag ttctggaccg accggctcgg gttctcccg gacttcgtgg aggacgactt 1920
60 cgccggtgtg gtccgggacg acgtgaccct gttcatcagc gcggtccagg accagggtgt 1980
61 gccggacaac accctggcct ggggtgtgggt gcgcggcctg gacgagctgt acgccgagt 2040
62 gtcggaggtc gtgtccacga acttccggga cgcctccggg ccggccatga ccgagatcgg 2100
63 cgagcagccg tgggggcggg agttcgccct gcgcgacccg gccggcaact gcgtgcactt 2160
64 cgtggccgag gagcaggact gacacgtgct acgagatttc gattccaccg ccgccttcta 2220
65 tgaaagggtg ggcttcggaa tcgttttccg ggacgccggc tggatgatcc tccagcgcg 2280
66 ggatctcatg ctggagttct tcgcccaccc caacttgttt attgcagctt ataatggtta 2340
67 caaataaagc aatagcatca caaatttcac aaataaagca tttttttcac tgcattctag 2400
68 ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tgtataccgt cgacctctag 2460
69 ctagagcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt atccgctcac 2520
70 aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctggggtg cctaagtagt 2580
71 gagctaactc acattaattg cgttgccgtc actgcccgtt ttccagtcgg gaaacctgtc 2640
72 gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattggg 2700
73 ctcttccgct tcctcgtcct ctgactcgtc gcgctcggtc gttcggctgc ggcgagcggt 2760
74 atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa 2820
75 gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 2880
76 gtttttccat aggtcccgcc ccctgacga gcatcacaaa aatcgacgct caagtcagag 2940
77 gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa gctccctcgt 3000
78 gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 3060
79 aagcgtggcg ctttctcaat gctcacgctg taggtatctc agttcgggtg aggtcgttcg 3120
80 ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 3180
81 taactatcgt cttgagtcca acccggttag acacgactta tcgccactgg cagcagccac 3240
82 tggtaacagg attagcagag cgaggtagtg aggcggtgct acagagttct tgaagtgggt 3300
83 gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt 3360
84 taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 3420
85 tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 3480
86 tttgatcttt tctacggggg ctgacgctca gtggaacgaa aactcacgtt aagggatttt 3540
87 ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 3600
88 taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 3660
89 tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt 3720
90 cgtgtagata actacgatac gggagggtt accatctggc ccagtgctg caatgatacc 3780
91 gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc 3840
92 cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta attggtgccc 3900
93 ggaagctaga gtaagtagtt cgccagttaa tagtttgccg aacgttggtg ccattgctac 3960
94 aggcacgtg gtgtcacgct cgtcgttttg tatggcttca ttcagctccg gttcccaacg 4020
95 atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct ccttcgggtc 4080
96 tccgatcgtt gtcagaagta agttggccgc agtggtatca ctcatggtta tggcagcact 4140
97 gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc 4200
98 aaccaagtca ttctgagaat agtgatgcg gcgaccgagt tgctcttgcc cggcgtcaat 4260
99 acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc 4320
100 ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaaccac 4380
101 tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa 4440
102 aacaggaagg caaatgccg caaaaaagg aataaggcg acacggaaat gttgaatact 4500
103 catactcttc ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg 4560

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,481

DATE: 01/27/2002

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

```

104 atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg 4620
105 aaaagtgcc cctgacgtc 4639
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 5
109 <212> TYPE: DNA
110 <213> ORGANISM: Artificial Sequence
112 <220> FEATURE:
113 <223> OTHER INFORMATION: single-stranded restriction fragment overhand
115 <400> SEQUENCE: 2
116 gcagg 5
118 <210> SEQ ID NO: 3
119 <211> LENGTH: 5
120 <212> TYPE: DNA
121 <213> ORGANISM: Artificial Sequence
123 <220> FEATURE:
124 <223> OTHER INFORMATION: single-stranded restriction fragment overhang
126 <400> SEQUENCE: 3
127 tgaga 5
129 <210> SEQ ID NO: 4
130 <211> LENGTH: 33
131 <212> TYPE: DNA
132 <213> ORGANISM: Artificial Sequence
134 <220> FEATURE:
135 <223> OTHER INFORMATION: recognition site
137 <221> NAME/KEY: misc_feature
138 <222> LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 22, 23, 24,
139 25, 26, 27, 28, 29, 30, 31, 32, 33
140 <223> OTHER INFORMATION: n = A,T,C or G
142 <400> SEQUENCE: 4
143 nnnnnnnnnn acnnngtay cnnnnnnnnn nnn 33
145 <210> SEQ ID NO: 5
146 <211> LENGTH: 33
147 <212> TYPE: DNA
148 <213> ORGANISM: Artificial Sequence
150 <220> FEATURE:
151 <223> OTHER INFORMATION: recognition site
153 <221> NAME/KEY: misc_feature
154 <222> LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 19,
155 20, 21, 27, 28, 29, 30, 31, 32, 33
156 <223> OTHER INFORMATION: n = A,T,C or G
158 <400> SEQUENCE: 5
159 nnnnnnnnnn nnnntggnn ncatrgnnnn nnn 33
161 <210> SEQ ID NO: 6
162 <211> LENGTH: 10
163 <212> TYPE: DNA
164 <213> ORGANISM: Artificial Sequence
166 <220> FEATURE:
167 <223> OTHER INFORMATION: modification fragment
169 <400> SEQUENCE: 6

```

RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/804,481

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

```

170 cacaacacaca 10
172 <210> SEQ ID NO: 7
173 <211> LENGTH: 12
174 <212> TYPE: DNA
175 <213> ORGANISM: Artificial Sequence
177 <220> FEATURE:
178 <223> OTHER INFORMATION: modification fragment
180 <400> SEQUENCE: 7
181 tccacaaaca ca 12
183 <210> SEQ ID NO: 8
184 <211> LENGTH: 15
185 <212> TYPE: DNA
186 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: modification fragment
191 <400> SEQUENCE: 8
192 tcgtccacaa acaca 15
194 <210> SEQ ID NO: 9
195 <211> LENGTH: 12
196 <212> TYPE: DNA
197 <213> ORGANISM: Artificial Sequence
199 <220> FEATURE:
200 <223> OTHER INFORMATION: modification fragment
202 <400> SEQUENCE: 9
203 cacaacacac ac 12
205 <210> SEQ ID NO: 10
206 <211> LENGTH: 10
207 <212> TYPE: DNA
208 <213> ORGANISM: Artificial Sequence
210 <220> FEATURE:
211 <223> OTHER INFORMATION: modification fragment
213 <400> SEQUENCE: 10
214 cacaacacacg 10
216 <210> SEQ ID NO: 11
217 <211> LENGTH: 59
218 <212> TYPE: DNA
219 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: vector construct
224 <400> SEQUENCE: 11
225 ggccaagat ctcaagggcc cataacatgt gtaccatcga ttgcagggga gataccatg 59

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/804,481

DATE: 01/27/2002

TIME: 15:27:50

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5